CENTRE DE RECHERCHE SUR LES TRANSPORTS
(C.R.T.)
CENTRE FOR RESEARCH ON TRANSPORTATION

UNIVERSITÉ DE MONTRÉAL
ÉCOLE DES HAUTES ÉTUDES COMMERCIALES
ÉCOLE POLYTECHNIQUE

CENTRE DE RECHERCHE SUR LES TRANSPORTS
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCCURSALE A
MONTRÉAL
CANADA H3C 3J7

TÉLÉPHONE: (514) 343-7575
TÉLEX: (21) 05-581325
EMAIL: crt@crt.umontreal.ca
TÉLÉCOPIEUR (FAX): (514) 343-7121

CRT-714

Forecasting Short-Term Demand for Empty Containers: A Case Study

by
Michel Gendreau[1]
Teodor Gabriel Crainic[1,2]
Pierre Dejax[3]
Hélène Steffan[1,3]

[1] Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale A, Montréal, Canada H3C 3J7.
[2] Département des sciences administratives, Université du Québec à Montréal, C.P. 6192, succursale A, Montréal, Québec H3C 4R2
[3] Laboratoire Economique, Industriel et Social, Ecole Centrale Paris, Chatenay-Malabry, France.

Centre de recherche sur les transports — Publication #714

September 1991

92-00718/

Dépôt légal – 4ème trimestre 1991

ABSTRACT

The purpose of the study described in this paper is to explore issues related to the modeling and estimation of short term demand of empty containers for subsequent movements on international shipping lines. The study is part of much larger research effort directed towards the development of models, methods and integrated planning tools to address typical problems related to the management of the land distribution and transportation of containers. The study is based on actual operational data coming from a large international maritime shipping company. The paper describes this information and the extensive manipulations required in order to obtain a suitable data set for statistical analyses. These included analyses both on the daily demand and on the weekly demand for various combinations of container category and customer aggregation. We conclude with a discussion of the difficulties and requirements associated with forecasting short-term demand of containers.

KEYWORDS Short Term Demand Analyses;
Time Series Analyses;
Land Container Distribution and Transportation;
Vehicle Allocation; Fleet Management.

RESUME

Le but de l'étude décrite dans cet article est d'explorer la problématique reliée à la modélisation et à l'estimation de la demande à court terme de conteneurs, dans le cadre du commerce international. Cette étude s'insère dans un programme de recherche dont le but est le développement de modèles, de méthodes et d'outils d'aide à la décision pour la résolution de problèmes liés à la gestion de la partie terrestre de la distribution et du transport de conteneurs. L'étude est basée sur des données d'opération provenant d'une companie de transport international maritime de conteneurs. L'article décrit les données disponibles, ainsi que le traitement requis par leur mise-en-forme en vue d'analyses statistiques. Ces analyses ont porté principalement sur la demande quotidienne et hebdomadaire pour diverses combinaisons de types de conteneurs et d'agrégation de clients. Nous concluons avec une discussion des difficultés et exigences associées à la prévision de la demande à court terme de conteneurs.

MOTS CLES Analyse de demande à court terme;
Séries chronologiques;
Transport terrestre de conteneurs;
Gestion de flotte; Allocation de véhicules.

I. INTRODUCTION

The purpose of this paper is to present a study aimed at the modeling and estimation of short term demand of empty containers within the land networks of international shipping lines. The study is part of much larger research effort directed towards the development of models, methods and integrated planning tools to address typical problems encountered by international shipping companies, which operate large scale maritime and land networks, with the management of their container fleet and of their land distribution and transportation operations.

The planning of these operations is an extremely complex activity, especially if one aims to simultaneously optimize the cost and service aspects of the company's operations in a competitive environment. Crainic, Gendreau and Dejax (1991) describe the various operations and planning issues involved, and present the methodological framework that we propose in order to address them. To facilitate the understanding of the context of the demand study, we now present a brief overview of operations and of the proposed methodology.

Arriving ships carry containers, which come in several sizes and types, loaded with imported goods, as well as empty containers returning from previous exports. Loaded containers are moved, by (multimodal) transportation modes, to their final destinations (import customers), while imported empty containers stay at the port are dispatched wherever they are needed for subsequent exports. Once unloaded, empty containers are up at the customer's site, and return either to the port of origin or to any other depot. Symmetrically, exporting customers require empty containers. Once loaded, containers are transported to the export port and are loaded on ships together with empty containers sent abroad to cope with the worldwide supply/demand imbalance. Note the importance of empty container movements. First, every commercial (profitable) movement of a loaded container generates almost automatically a non-profitable empty container movement. Second, significant regional imbalances between imports and exports result in empty containers being moved over relatively long distances, usually directly between depots. Thus, up to 40% of the land container traffic is made up of empty movements, which represent a significant portion of the total system cost (Dejax et al., 1987; Dejax, Crainic and Delorme, 1988).

The methodology proposed to improve the management of empty container movements is based on an integrated multilevel approach which reflects the observed hierarchy in the decision process and the flow of information. Its first main component is a strategic/tactical model, formulated as a multimode, multicommodity location/distribution problem with inter-depot balancing requirements (Crainic, Dejax and Delorme, 1989; see Crainic, Gendreau and Dejax, 1991, for references concerning algorithms developed for this model). The output of this model consists of the set of depots to be used for the duration of the planning horizon (several months to a year), the customer to depot allocation rules, and the main interdepot empty container balancing flows.

The second component of our method corresponds to the level of the operational (day-to-day) planning of the company's activities. At this level, one has to make sure that demand is satisfied and that the most effective routes and means of transportation are selected and used. Two models are used: an *empty container allocation* model and an *empty and loaded container routing* model. The allocation model aims to determine the "best" distribution of empty containers that satisfies known and forecasted customer demands at lowest total system cost; it is formulated as a stochastic, dynamic network model (Crainic, Gendreau and Dejax, 1990). The routing model strives to minimize the overall transportation cost of the loaded and empty containers from their origin to their destination, while insuring on-time delivery (Crainic et al., 1991).

Demand data, especially short-term forecasts, is essential for these operational models since, besides regular customers with sufficiently well known behaviour (and possible long term contracts), a significant part of the total service requests come from irregular customers, which order rarely, or at irregular intervals, or both. Yet, to our knowledge, there does not exist any model with which one could estimate the short term (daily, weekly) container demand. Hence, we have initiated an empirical investigation aimed at determining possible formulae for the average demand, and the associated probability distributions. Ultimately, we look for models which may help forecast demand by container type and by geographical zone, for a very short horizon. This paper gives an account of the first phase of this study, dedicated to the exploration of the available information, the construction of a reliable data set, and its statistical analysis. The plan of the paper parallels this sequence.

II. THE AVAILABLE DATA

The initial data available for this study comes from a company which operates worldwide maritime shipping lines to and from some twenty major European ports, while servicing customers throughout most of Western Europe. The data corresponds to the land distribution and transportation operations of the company for the year 1986. We concentrate our study on a network which covers France, parts of Germany (corresponding to the former Federal Republic), The Netherlands, Belgium and Luxemburg. This area corresponds to about 80% of the total container movements the company performed in Europe in that particular year.

The 224,374 records of the data file stand for as many land movements of one or several loaded or empty containers. Amongst the various pieces of information recorded for each movement, the most significant for the demand study are:

- *Transportation mode*: identifies the actual mode used for transportation and the type of movement. The two main transportation means are private trucking and railways, including multimodal truck-rail combinations. Movements may be import or export, loaded or empty, commercial or technical (balancing flows, movements of damaged or rented containers, etc.),

- *Commercial mode*: specifies the type of contract between company and customer. Two modes are prevalent: 1) *carrier haulage* (70% of the performed movements) where the company manages, and pays for, the transportation of the containers; and 2) *merchant haulage* (some 20% of the performed movements) where the customer is entirely in charge of moving the containers and only the loaded movements are registered (with zero cost) in the company's records. The remaining traffic moves under *mixed mode* agreements, where the company and the customer are each responsible for a part of the trip.

- *Container category*. In 1986, the company used containers of 20 different categories (see Table 1). The specific container category used for a given movement is identified by a combination of *size* and *type* characteristics. Note that, with the exception of the "bulk" and $9m^3$ containers, the company used exclusively 20' and 40' containers (in Europe, a truck is 20' or 40' long, while a rail car measures 60'). Distinctions among these containers are then induced by the particular usage (shipping line or operational

characteristics) they are intended for.

Size	Type	Designation	Particularities
05	00	$9m^3$ containers	one 20' = 3 $9m^3$
20	00	20' general	All heights
22	00	20' general	8'6 height (Zeebruge)
43	00	40' general	
20	51	20' open top	
40	51	40' open top	
20	60	20' flat	
45	00	40' flat	
20	40	20' isotherm	Not allowed on French Caribbean lines
42	40	40' isotherm	Not allowed on French Caribbean lines
22	49	20' isotherm	Dedicated to French Caribbean lines
42	49	40' isotherm	Dedicated to French Caribbean lines
22	32	20' refrigerated	
43	32	40' refrigerated	
20	72	20' tanker	
41	CT	40' tanker	
49	RW	ACL trailer	ACL shipping line
VR		bulk	
NU		empty trailer	Truck movement without a container
22	80	20' special bulk	
22	02	20' open side	

Table 1. Description of Container Categories

Several manipulations of this initial information had to be performed in order to obtain a data set suitable for demand analyses (see Steffan, 1989 for details). After cleaning up the file, we identified the origins and destinations of movements as customers or depots. We then determined a *customer (depot) zone* for each customer (depot) by selecting some 300 points where container traffic is most intense, and then by building around these points zones which are consistent both geographically and commercially. Thus, since no aggregation of records has been performed, we insure that, in most cases, a meaningful number of occurrences (observations) for each possible movement type is available for statistical analyses. The next step consists in the estimation of the daily supply and demand of empty containers, for each customer zone and container category.

Two approaches may be used to estimate the empty container supply and demand. The first approach counts the *empty* containers which arrive (the empty customer demand) at and

leave (the empty customer supply) from each customer zone. This approach cannot be applied to movements performed under the merchant haulage commercial arrangement. The second method is based on the fact that each time a customer ships out a loaded container, the company had first to deliver an empty one and, symmetrically, each loaded container delivered to a customer has, later on, to be picked up and moved away empty. Thus, the second approach counts the *loaded* containers with leave (the empty customer demand) and arrive (the empty customer supply) at each customer site.

Category	Carrier Haulage and Mixed Mode				Merchant Haulage	
	Empty		Loaded		Loaded	
	Demand	Supply	Supply	Demand	Demand	Supply
0500	29	8	8	27	12	505
2000	14183	6414	11361	18302	4674	13478
2040	410	2478	2920	416	628	193
2051	1004	482	520	1024	19	478
2060	223	90	93	242	80	439
2072	351	438	440	352	303	149
2200	3102	2732	3463	2931	549	1803
2232	1066	836	855	1080	378	1502
2249	6229	4568	6734	7181	14446	4320
2280	9	0	0	0	0	0
4051	1029	174	208	1013	10	157
41CT	3	11	11	3	1	0
4240	120	38	47	114	18	5
4249	1564	974	1346	1649	1950	746
4300	6465	2666	4117	8132	1137	2589
4332	1214	876	1015	1348	31	303
4500	581	56	36	560	7	1021
49RW	115	40	43	117	0	0

Table 2. Supply/Demand Evaluations

Table 2 displays the results of the two methods for our data set. One notes that both methods yield approximately the same figures. Yet, the total supply/demand is higher when the *loaded* method is used, which indicates that not all empty movements have been originally recorded in the company's data base. Consequently, we based our demand study on the daily estimations obtained by using this approach.

To conclude this section, we want to mention two sources of potential difficulties in the analysis of data and the interpretation of results. First, as emphasized by the yearly figures displayed in Table 2, not all container types are equally used. In fact, some of them are very rarely called for. Consequently, one year data may not be sufficient to obtain statistically significant results for some container types. Second, it should be clear that the available data does not reflect demand but rather the actual operations: the performed movements of loaded containers and some of the empty ones. Hence, the date of the request relative to a given movement is not recorded in the available data, and we miss the refused demand, as well as the container substitutions the company had to make in order to satisfy demand when the particular container category requested was not available. Furthermore, it has not been possible to establish exactly the meaning of the temporal information recorded with each movement. (At that time, the company was contracting out all its computer related operations, and was not using, nor validating in detail, its past operational data.) Based on our discussions with the company's management, we assumed that it indicates the date when the movement took place. This hypothesis is somehow corroborated by the statistical analyses we performed. However, it may also indicate, at least for some movements, the date when the operation has been recorded, thus introducing a certain level of uncertainty in our analysis.

In spite of these problems, it was decided to perform the analyses on the movement data, interpreting movements as demand, since this is the only information available. Furthermore, this is the type of data which is very likely to continue to be available, since the planned management information system of the company was still meant to record the actual performed movements.

III. STATISTICAL ANALYSES

In spite of the aforementioned uncertainties regarding the exact timing of customer requests for empty containers, a series of statistical analyses were performed on the final demand data set. These included analyses both on the daily demand and on the weekly demand for various combinations of container category (also called, for the sake of brevity, *product*, in this section) and customer aggregation.

3.1 Daily demand analyses

These analyses were motivated by the fact that, in many fields of freight transportation, demand follows weekly cyclic patterns (see, for instance, Beaujon and Turnquist, 1991). More specifically, we were interested in determining whether a model of the form

$$D^p_{dw} = D^p_w \, \theta^p_d$$

with

D^p_{dw} : demand for product p on day d of week w,

D^p_w : demand for product p on week w,

θ^p_d : day-of-the-week adjustment factor for product p and day d,

could be fitted to the observed data.

Plotting demands for every day of the week over the 52 weeks of the year did not yield any clear pattern, whether one considered aggregate demands for a product or demands for a specific (container category, customer zone) combination, apart from the fact that observed values for Saturdays and Sundays were much lower.

To further investigate the impact of the day-of-the-week on the daily demands, linear regressions with dummy variables associated to the days of the week were performed on the data streams with the largest observed values, i.e., the aggregate demands per product and the demands for 24 (product, zone) pairs. In these regressions, the Saturday and Sunday values of each week were added together yielding 6 observations per week for each data stream. The general form of the resulting regression equations was thus:

$$D = aX_1 + bX_2 + cX_3 + dX_4 + eX_5 + f$$

where D is the daily demand for a given product or (product, zone) pair, X_i equals 1 if the demand occurs on the i-th day of the week (and 0 otherwise), and a, b, c, d, e and f are the regression coefficients to be estimated.

Category	R^2	F-ratio	Probability
0500			
2000	41%	32.49	0.0000
2040	10%	7.18	0.0000
2051	4%	13.13	0.0003
2060	1%	4.63	0.0323
2072	14%	9.87	0.0000
2200	29%	24.86	0.0000
2232	9%	15.43	0.0000
2249	55%	76.54	0.0000
4051	16%	12.05	0.0000
4240	2%	7.21	0.0076
4249	39%	38.57	0.0000
4300	47%	53.61	0.0000
4332	20%	19.02	0.0000
4500	11%	7.89	0.0000
49RW	7%	7.70	0.0001

Table 3. Regression Results on Unnormalized Aggregate Data

The results of these regressions are summarized in Tables 3 and 4. In Table 3, it must be pointed out that for product 0500 none of the independent variables X_i was sufficiently correlated with the dependent variable D to pass the tolerance tests of the stepwise regression method; in this case, the regression is thus not significant. For the other products, the F-ratios indicate that the regressions are globally significant (the last column of the table gives the probability that the regression as a whole is not significant), but the R^2 coefficients, which correspond to the proportion of the total variation which is explained by the model, are not very large. For the (product, zone) pairs (Table 4), though almost all regressions can be considered as significant, the R^2 coefficients are even smaller, which indicates a fairly poor fit.

Obviously, seasonal patterns could occur; if such was the case, they would probably be responsible for the large amount of unexplained variation. To account for such patterns, new data were created by dividing daily demands by weekly demands (the new observations thus represent the proportion of weekly demand which takes place on a given day). The

Category	Zone	R^2	F-ratio	Probability
2000	105	8%	5.54	0.0001
2000	107	13%	9.32	0.0000
2000	140	9%	6.47	0.0001
2000	154	6%	3.55	0.0039
2000	155	6%	3.81	0.0023
2000	210	8%	5.12	0.0002
2000	220	4%	2.27	0.0472
2000	222	4%	2.80	0.0268
2000	224	6%	3.85	0.0022
2000	237	17%	12.54	0.0000
2000	239	7%	4.84	0.0003
2232	208	12%	7.80	0.0000
2249	9	2%	1.60	0.1753
2249	140	23%	17.64	0.0000
2249	172	17%	12.49	0.0000
2249	192	8%	5.34	0.0001
2249	237	20%	15.27	0.0000
2249	274	33%	29.04	0.0000
2249	275	11%	7.59	0.0000
4300	107	7%	4.84	0.0003
4300	140	15%	10.69	0.0000
4300	208	21%	16.32	0.0000
4300	210	13%	9.13	0.0000
4300	239	2%	0.93	0.4586

Table 4. Regression Results on Unnormalized Disaggregate Data

corresponding regression equation is then:

$$\frac{D}{D_w} = a'X_1 + b'X_2 + c'X_3 + d'X_4 + e'$$

where D_w is the weekly demand and D, X_1, ..., X_4 are defined as before. (Note that one independent variable had to be eliminated, since the observed values for every week now sum to 1).

The results of these regressions for the global demands per product are reported in Table 5. Surprisingly, the normalization of the daily demands does not produce much larger R^2 values overall and there are now two container categories for which the regressions are not significant (0500 and 4240).

Category	R^2	F-ratio	Probability
0500			
2000	42%	43.66	0.0000
2040	11%	7.42	0.0000
2051	16%	11.86	0.0000
2060	4%	5.89	0.0031
2072	13%	9.32	0.0000
2200	33%	29.99	0.0000
2232	17%	12.29	0.0000
2249	54%	73.41	0.0000
4051	15%	18.58	0.0000
4240			
4249	39%	39.84	0.0000
4300	50%	61.46	0.0000
4332	34%	31.30	0.0000
4500	12%	8.65	0.0000
49RW	8%	13.06	0.0000

Table 5. Regression Results on Normalized Aggregate Data

It was pointed out earlier that demands on Saturdays and Sundays are, not unexpectedly, much lower. Since this could have had an impact on the previous analyses, it was decided to discard these observations and to perform for the global demands per container category a new series of regressions similar to the first one (i.e., with unnormalized values). These regressions yielded much lower R^2 coefficients (for instance, for product 2000 the R^2 goes down from 41% to 18% and for product 2249 from 55% to 17%); this confirmed that, apart from the difference between working days and weekends, the day-of-the-week effect is not very important. This led us to abandon the model initially proposed and to focus our attention on weekly demands.

3.2 Weekly demand analyses

While our analyses of daily demands were aimed at the identification of repetitive cyclic patterns, a different approach had to be used with regards to weekly demands. For one thing, the available data (1 year) is clearly insufficient to provide any insight relative to possible seasonal patterns in any of the data streams. Furthermore, preliminary analyses of the aggregate demands per container category indicated that, if seasonal patterns were indeed present, these

would be different from one product to another, thus making it useless to try to identify a general pattern for all categories. One the other hand, one may wonder about the possibility that the observed variations in the demands simply correspond to plain randomness in the underlying processes. To test this hypothesis, we first tried to fit Normal and Poisson distributions to the demands per container category. The parameters used for the postulated distributions were the sample mean and standard deviation in case of the Normal, and sample mean for the Poisson. The goodness-of-fit test was the Kolmogorov-Smirnov test which compares the observed cumulative distribution to the postulated cumulative distribution on the basis of the MEAD ("most extreme absolute difference") between them. In fact, in this test, the MEAD is transformed into a standard Normal statistic (Z value) from which the probability that the observed data follow the postulated distribution can be derived.

Category	Mean	Standard Deviation	MEAD	Kolmogorov Smirnov Z	Probability
0500	0.53	1.03	0.364	2.597	0.000
2000	354.10	58.30	0.053	0.381	0.999
2040	8.08	5.11	0.153	1.094	0.182
2051	19.75	11.46	0.159	1.136	0.151
2060	4.51	4.91	0.205	1.467	0.027
2072	6.84	3.82	0.106	0.760	0.610
2200	57.18	20.90	0.066	0.469	0.980
2232	20.84	7.83	0.139	0.993	0.277
2249	138.69	23.12	0.076	0.541	0.932
4051	19.67	7.50	0.099	0.709	0.696
4240	2.18	2.46	0.215	1.534	0.018
4249	32.08	8.01	0.143	1.019	0.250
4300	158.27	27.57	0.069	0.494	0.968
4332	26.18	13.79	0.176	1.255	0.086
4500	10.82	6.29	0.116	0.830	0.496
49RW	2.27	2.59	0.218	1.556	0.016

Table 6. Kolmogorov-Smirnov Test: Normal Distribution

The results of these tests (see Tables 6 and 7) indicate that the Normal distribution provides a good fit for four products (2000, 2200, 2249 and 4300) and a reasonable one for two (2072 and 4051), while the demands for container categories 0500 and 4249 are approximately Poisson. It is interesting to note that the four container categories for which a good fit was obtained with the Normal distribution are those with the largest sample means; there is

thus a high level of aggregation of individual demands for these products and, therefore, this result could be expected.

It should be pointed out that the Kolmogorov-Smirnov test considers all observations as a static sample neglecting any serial autocorrelation within the data streams. This is quite important since the presence of significant autocorrelations would invalidate the assumption of plain randomness. When we computed the autocorrelations and the partial autocorrelations for the 16 series, we found out that for 9 of these (products 2000, 2051, 2060, 2072, 2232, 2249, 4051, 4240 and 49RW), there were no significant autocorrelations. It is thus possible to conclude that the series corresponding to container categories 2000, 2072, 2249 and 4051 are indeed "white noise", i.e., sequences of independent identically distributed Normal random variables.

Category	Mean	MEAD	Kolmogorov Smirnov Z	Probability
0500	0.53	0.078	0.555	0.918
2000	354.10	0.290	2.072	0.000
2040	8.08	0.136	0.973	0.300
2051	19.75	0.231	1.648	0.009
2060	4.51	0.253	1.808	0.003
2072	6.84	0.143	1.022	0.247
2200	57.18	0.263	1.878	0.002
2232	20.84	0.182	1.298	0.069
2249	138.69	0.186	1.330	0.058
4051	19.67	0.168	1.203	0.111
4240	2.18	0.169	1.207	0.108
4249	32.08	0.078	0.558	0.915
4300	158.27	0.226	1.617	0.011
4332	26.18	0.358	2.569	0.000
4500	10.82	0.198	1.414	0.037
49RW	2.27	0.193	1.376	0.045

Table 7. Kolmogorov-Smirnov Test: Poisson Distribution

For the 7 other container categories, further analysis was required. This was done by using the well-known Box-Jenkins method which allows one to characterize sequential dependencies in time series through auto-regressive integrated moving average (ARIMA) models (Box and Jenkins, 1976). For two of the container categories (2040 and 4300), it was

impossible to identify satisfactory models. For the five remaining products, models were identified, but they all displayed a very high residual variance to mean ratio, which would make them almost useless in practice for predictive purposes (see Steffan, 1989, for more details).

We also decided to apply the Box-Jenkins method to the disaggregate series, i.e., the series for the (container category, customer zone) pairs. For practical reasons, this analysis was restricted to the 24 pairs with annual demand larger than 300 containers. For these disaggregate series, the Box-Jenkins method yielded the following results (an exhaustive report on these analyses can be found in Steffan, 1989):

- for 10 series, no ARIMA model could be identified;
- 4 series displayed no significant autocorrelations and it was possible to fit a Normal distribution for three of them and a Poisson for the last one;
- ARIMA models were identified and fitted for the other 10 series; however, these models were quite different from one another and, in most cases, the residual variances were large.

Overall, the results obtained for the disaggregate series are just a confirmation of what could be suspected after the analysis of the aggregate series: the processes underlying the demand for empty containers are extremely complex and their combined effect cannot be easily characterized in statistical terms, except when the level of aggregation is very high (that would be the case of the global demands for products 2000, 2072, 2249 and 4051).

3.3 Discussion

Several general remarks are in order with regard to the analyses that were performed. First, with only one year of data, it is not possible to detect any overall trend in the demands, nor to identify any seasonal patterns which one could expect. Second, container categories with low traffic display high coefficients of variation, even for the weekly demands. This may suggest the use of different probability distributions, such as the Negative Binomial or the Gamma, to represent the demand in these cases. Third, some external factors have an important effect on the observed demands. For one thing, demands for empties are certainly driven to some extent

by ship schedules, since a request for a container probably occurs a few days before the departure date of the ship on which it will sail. With regard to the daily variations in demand, one must also take into account legal holidays and "traditional" vacation periods (such as most of July and August in most of Western Europe). When we plotted the daily demands, we were able to easily pick up on the graphs the dates of all major holidays: in each case, there was a very significant dip in demand. Given the large number of such holidays in the countries covered by the study, this in itself may have been sufficient to "throw off" the time series analyses.

IV. CONCLUSIONS

As stated in the Introduction, the objective of the empirical study described in this paper was to gain some insight into the short term demand process for empty containers and, if possible, to derive demand forecasting models.

We were successful in constructing a data set on demand for empties by using recorded information on loaded container movements. This data set, though not perfect, was adequate enough to allow for statistical analyses of demand.

The statistical analyses that were performed have clearly shown that predicting short-term demand for empty containers is extremely difficult. This rejoins the conclusions of similar studies performed for the rail mode (Minger, Williams and Hargrove, 1975; Minger and Hargrove, 1977). However, when the traffic is consistently large, we were able to fit probability distributions for the demand of specific container categories. These distributions can certainly be directly used in short-term planning models.

The situation is different for container categories with low traffic. For these, the available data was not sufficient to estimate distributions by using straightforward statistical techniques. For good predictions to be made, one needs more than one year of data, and one must take into account all elements which may have a significant impact on the demand process in the specific case under study: ship schedules, holidays, vacations, substitution rules, etc. Note however, that the lack of reliable demand distributions for these container categories is not a critical issue in short-term planning. This is due to the fact that conservative estimates of buffer stocks at depots can be derived from the means and variances of observed demand. Moreover, given the low level of demands (and stocks) for these containers, the inventory costs associated with overestimating demand form a very small portion of the total system costs.

ACKNOWLEDGEMENTS

The authors are grateful to the Canadian *National Science and Engineering Research Council*, to the *Fonds F.C.A.R.* of the Province of Québec, and to the *Programme de coopération France – Québec*, who have provided the financial support for this research.

REFERENCES

Beaujon G.J. and Turnquist M.A. (1988), *A Model for Fleet Sizing and Vehicle Allocation*, **Transportation Science** 25(1), 19-45.

Box G.E.P. and Jenkins G.M. (1976), *Time Series Analysis: Forecasting and Control*, Holden-Day.

Crainic T.G., Gendreau M. and Dejax P.J. (1991), *Modelling the Container Fleet Management Problem Using a Stochastic Dynamic Approach, Operational Research'90*, (H.E. Bradley Editor), Pergamon Press, 473-486.

Crainic T.G., Dejax P.J., Gendreau M. and Benamar F. (1991), *Multimodal Multiproduct Combined Vehicle Routing - Traffic Itinerary Generation*, **ORSA/TIMS Joint National Meeting**, Anaheim, November 3-6.

Crainic T.G., Gendreau M. and Dejax P.J. (1990), *A Dynamic Stochastic Model for the Allocation of Empty Containers*, publication #713, Centre de recherche sur les transports, Université Montréal, to appear **Operations Research**.

Crainic T.G., Dejax P.J. and Delorme L. (1989), *Models for Multimode Multicommodity Location Problems with Interdepot Balancing Requirements*, **Annals of Operations Research** 18, 279-302.

Dejax P., Crainic T.G., Delorme L., Bloise M., de Tocqueville E. and Hodgson J. (1987), *Planification tactique du transport terrestre des conteneurs vides*, Research report for an International Container Shipping Company, three volumes, Ecole Centrale de Paris, Paris, France.

Dejax P.J., Crainic T.G. and Delorme L. (1988), *Strategic/Tactic Planning of Container Land Transportation Systems*, Joint Meeting **TIMS XXVIII/EURO IX**, Paris, 6-8 juillet.

Minger W.K. and Hargrove M.B. (1977), *Freight Car Demand Information and Forecasting Research Project. Phase II*, Report no. FRA-OPPD-77/5, Association of American Railroads, Washington, D.C.

Minger W.K., Williams D.J. and Hargrove M.B. (1975), *Freight Car Demand Information and Forecasting Research Project. Phase I*, Report no. FRA-OPPD-76/9, Association of American Railroads, Washington, D.C.

Steffan H. (1989), *Etude de la demande de conteneurs vides a court terme*, Publication # 651, Centre de recherche sur les transports, Université de Montréal.